BU-336-M Feb 1970 p2/

A LOWER BOUND FOR THE NUMBER OF SINGULAR SATURATED MAIN EFFECT PLANS OF AN S^m FACTORIAL*

by

B. L. Raktoe and W. T. Federer

University of Guelph and Cornell University

0. SUMMARY

A fraction consisting of $m(s-1)^{+1}$ observations taken at $m(s-1)^{+1}$ treatment combinations of an s^m factorial with the aim to estimate the mean and the $m(s-1)$ main effect single degree of freedom parameters is called a saturated main effect plan. If the design matrix of such a fraction is singular, then the fraction is called a singular saturated main effect plan. This paper presents a lower bound on the number of singular saturated main effect plans and, conversely, an upper bound on the number of nonsingular saturated main effect plans.

1. INTRODUCTION

Consider the full replicate of an s^m factorial (s is a prime or a power of a prime), then it is well known (e.g. see Kempthorne [1952]), that the s^m treatment combinations form the points of the finite Euclidean geometry EG(m,s) over the field GF(s) and that the $(s^m - 1)/(s-1)$ effects are in 1:1 correspondence with the points of the finite projective geometry PG(m - 1, s). Each effect of the s^m factorial represents (s - 1) degrees of freedom, which in fact means that every

*Research partially supported by N.R.C. Grant #A07204 and Public Health Research Grant GM-05900.

point $(x_1, x_2, \ldots, x_m)$ of PG(m-1, s) represents the class $\rho(x_1, x_2, \ldots, x_m)$, where ρ is a non-zero element of GF(s). The elements of a class can then be taken to depict a set of (s - 1) single degree of freedom parameters. If we adjoin the mean $\mu = (A^0\ B^0\ \ldots\ M^0)$ to all single degree of freedom parameters, then it is readily seen that the s^m single degree of freedom parameters are in 1:1 correspondence with the points of EG(m,s). Hence we have explicitly that the set of treatment combinations $\{(z_1, z_2, \ldots, z_m), z_i \in GF(s)\}$ is in a 1:1 correspondence with the set of single degree of freedom parameters $\{(A^{z_1}\ B^{z_2}\ \ldots\ M^{z_m}), (z_1, z_2, \ldots, z_m) \in EG(m,s)\}$. This means that if we wish to discuss properties of both the treatment combinations and the single degree of freedom parameters we may limit ourselves to the finite Euclidean geometry EG(m,s).

If we have a full replicate of an s^m factorial, then the usual linear model tying up the observations and the single degree of freedom parameters (assuming that the factors have quantitative levels) is:

$$E[Y] = X\beta \qquad (1.1)$$

where: Y is an $s^m \times 1$ vector of observations, each component of which is taken at a treatment combination, X is an $s^m \times s^m$ matrix such that $X'X = \text{diagonal } (d_1, d_2, \ldots, d_N)$, $N = s^m$, and β is an $s^m \times 1$ vector of single degree of freedom parameters as described earlier.

<u>DEFINITION 1.1.</u> Define an element $A^{z_1}\ B^{z_2}\ \ldots\ M^{z_m}$ of β to be a main effect single degree of freedom parameter if the superscript $(z_1, z_2, \ldots, z_m)$ has exactly one non-zero coordinate.

DEFINITION 1.2. A plan consisting of $m(s - 1) + 1$ observations to estimate the mean μ and the $m(s - 1)$ main effect single degree of freedom parameters is termed a saturated main effect plan.

Now, if the $[m(s - 1) + 1]$ - vector of observations is denoted by Y_1, then we know that the normal equations for a saturated main effect plan is:

$$X'_{11} X_{11} \hat{\beta}_1 = X'_{11} Y_1 \tag{1.2}$$

where X_{11} is an $[m(s - 1) + 1] \times [m(s - 1) + 1]$ matrix simply read off from X of equation (1.1) and $\hat{\beta}_1$ is the least squares estimator of β_1, which is the $[m(s - 1) + 1]$ - vector with $\mu = A^0 B^0 \ldots M^0$ as its first element and the rest of the elements being main effect single degree of freedom parameters.

DEFINITION 1.3. Following Banerjee and Federer [1966] we define a saturated main effect plan to be singular if rank $[X_{11}] < m(s - 1) + 1$ and nonsingular if rank $[X_{11}] = m(s - 1) + 1$.

The question naturally arises as to which type of plans will yield singular saturated main effect plans or equivalently which $[m(s - 1) + 1]$ - subsets of EG(m,s) will give rise to singular X_{11} matrices. Also, if singular saturated main effect plans exist, then we wish to know how many of the

$$\binom{s^m}{m(s - 1) + 1} \tag{1.3}$$

possible plans are singular. These two questions can be considered as the general problem of singular (or nonsingular if the complementary problem is considered) saturated main effect plans.

The aim of this paper is to discuss a sub-class of singular saturated main effect plans, which lends itself to be treated geometrically and is in agreement with the theory of confounding. This then will lead us to a lower bound on the number of singular saturated main effect plans (and an upper bound on the number of nonsingular plans).

2. EXISTENCE AND ENUMERATION OF SINGULAR SATURATED MAIN EFFECT PLANS

Consider an incomplete block design consisting of s^m treatments in s^k blocks of s^{m-k} plots each, then it is well known (see for example Kempthorne [1952]) that the construction of such a design can always be done by providing a confounding scheme such that $(s^k - 1)/(s - 1)$ effects are completely confounded with the blocks. This statement of course is equivalent to providing a $(k - 1)$ - flat of $PG(m - 1,s)$, which in turn is equivalent to the exhibition of its set of k generators.

~~Suppose now that~~ If $G = \{A^{x_{11}} B^{x_{12}} \ldots M^{x_{1m}}, A^{x_{21}} B^{x_{22}} \ldots M^{x_{2m}}, \ldots, A^{x_{k1}} B^{x_{k2}} \ldots M^{x_{km}}\}$ is a set of k generators of a particular confounding scheme or equivalently of a particular $(k - 1)$ - flat of $PG(m - 1,s)$, then the following definition ~~will be needed in the sequel:~~ describe the particular set of treatment combinations in each block:

<u>DEFINITION 2.1.</u> The $(\alpha_1, \alpha_2, \ldots, \alpha_k)$ - th level of a set

of k generators, denoted by the symbol,

$$\{A^{x_{11}}B^{x_{12}}\ldots M^{x_{1m}}, A^{x_{21}}B^{x_{21}}\ldots M^{x_{2m}}, \ldots, A^{x_{k1}}B^{x_{k2}}\ldots M^{x_{km}}\}_{(\alpha_1,\alpha_2,\ldots,\alpha_k)} \tag{2.1}$$

will be defined to depict a set of s^{m-k} treatment combinations $\{(z_1, z_2, \ldots, z_m)\}$ in EG(m,s), satisfying the consistent and independent set of equations:

l.c.

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ x_{k1} & x_{k2} & \cdots & x_{km} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \cdot \\ z_m \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \cdot \\ \alpha_k \end{bmatrix}, \tag{2.2}$$

where $(\alpha_1, \alpha_2, \ldots, \alpha_k)$ is a given set of k elements from GF(s).

#7 Note that the solutions to (2.2) form a (m - k) - flat of EG(m,s). Also note, that for a fixed set of k generators of a confounding scheme we have precisely s^k fractional replicates of order s^{m-k}, since each of the α_i's can be chosen in s ways from GF(s). A particular fractional replicate of this type can then be denoted by:

$$I = \{A^{x_{11}}B^{x_{12}}\ldots M^{x_{1m}}, A^{x_{21}}B^{x_{22}}\ldots M^{x_{2m}}, \ldots, A^{x_{k1}}B^{x_{k2}}\ldots M^{x_{km}}\}_{(\alpha_1,\alpha_2,\ldots,\alpha_m)}$$

k

with the usual meaning that the mean is completely confounded with $(s^k - 1)/(s - 1)$ effects, generated by the k generators within the braces.

<u>DEFINITION 2.2</u>. Define k to be the largest positive integer such that we may select m(s - 1) + 1 treatment combinations

(i.e. a saturated main effect plan) from among the s^{m-k} combinations of the fraction (2.3). This definition of k implies the following inequalities for given m and s:

$$m(s-1)+1 \le s^{m-k} \tag{2.4}$$

$$\text{i.e.} \quad m \le (s^{m-k}-1)/(s-1) \tag{2.5}$$

If we denote a saturated main effect plan of $m(s-1)+1$ treatment combinations selected from among s^{m-k} treatment combinations where k satisfies (2.4) or (2.5), by D, then the following can be easily verified:

THEOREM 2.1. <u>The number of plans of type D is given by</u>

$$N[m-1,\ k-1,s] = \phi[m-1,k-1,s] \cdot s^{k} \cdot \binom{s^{m-k}}{m(s-1)+1} \tag{2.6}$$

<u>where</u> $\phi[m-1,k-1,s]$ <u>is the number of</u> $(k-1)$ - <u>flats in</u> $PG(m-1,s)$ <u>explicitly given by</u> (see Mann [1949]):

$$\phi[m-1,k-1,s] = \Big[\prod_{i=0}^{k-1} (s^{m-i}-1)\Big]\Big/\Big[\prod_{i=0}^{k-1} (s^{k-i}-1)\Big] \tag{2.7}$$

Of course, the number given by (2.6) is less than the number $T(m,s)$ of saturated plans selected in an unrestricted manner, i.e.

$$N[m-1,k-1,s] < \binom{s^{m}}{m(s-1)+1} = T(m,s) \tag{2.8}$$

Now, let $T[m,s,0]$ denote the number of all singular saturated main effect plans, i.e. those plans from among the $T(m,s)$ plans which lead to singular X_{11} matrices in our setting (1.2), i.e., those plans for which $|X_{11}| = 0$. (The number $T[m,s,0]$ is ~~not~~

known for ~~all~~ very few s^m factorials and this problem and some related ones are currently under study). Let N[m-1,k-1,s,0] denote the total number of singular saturated main effect plans of type D as described earlier, then our intention is to determine N[m-1,k-1,s,0] and to show that this number is a lower-bound to T[m,s,0]. Naturally we have to discuss first the existence of singular saturated main effect plans. The following theorem establishes the existence of singular saturated main effect plans:

THEOREM 2.2. <u>If a main effect or a two-factor interaction is completely confounded with the mean, then the fraction leads to a singular saturated main effect plan.</u>

PROOF: Let k satisfy the inequality (2.4) or (2.5) and let m(s - 1) + 1 treatment combinations be selected from the fraction

$$I = \{A^{x_{11}}B^{x_{12}}\ldots M^{x_{1m}},\ldots,A,\ldots,A^{x_{k1}}B^{x_{k2}}\ldots M^{x_{km}}\}_{(\alpha_1,\alpha_2,\ldots\alpha_m)} \tag{2.9}$$

Here we have for simplicity, but without loss of generality, chosen the main effect A (A represents, as pointed out in section 1, a set of (s - 1) main effect single degree of freedom parameters). From (2.9) it follows immediately that in the design matrix X_{11} of (1.2) the columns corresponding to $A^{u_1=1}, A^{u_2},\ldots,A^{u_{s-1}}$ will have columns of the form $c_1\mathbf{1}$, $c_2\mathbf{1}$, $\ldots,c_{s-1}\mathbf{1}$, where the u_i's are non-zero elements of GF(s), ~~and~~ the c_i's are integers, and ~~finally~~ $\mathbf{1}$ is an [m(s - 1) + 1] - column vector of +1's. Hence it follows that rank $[X_{11}]<m(s-1)+1$,

i.e. X_{11} is singular. Similarly if we select $m(s - 1) + 1$ combinations from the fraction:

$$I = \{A^{x_{11}}B^{x_{12}}\ldots M^{x_{1m}},\ldots,AB,\ldots,A^{x_{k1}}B^{x_{k2}}\ldots M^{x_{km}}\}_{(\alpha_1,\alpha_2,\ldots,\alpha_m)} \qquad (2.10)$$

where without loss of generality we have taken the two factor interaction AB, then by the usual group theoretic multiplication we obtain that A is completely confounded with B. This then as above immediately implies the singularity of X_{11}.

This theorem implies that the class of all singular saturated main effect plans of type D is completely characterized by the set of confounding schemes in which either a main effect or two-factor interaction is confounded, this set arising from the consideration of an incomplete block design with s^m treatments in s^k blocks of s^{m-k} plots each, where k satisfies either (2.4) or (2.5). The number $N[m-1, k-1, s, 0]$ of such plans can be determined from the following theorem.

THEOREM 2.3. <u>Let k be the largest integer such that for given m and s the inequality</u> (2.4) <u>or</u> (2.5) <u>is satisfied and let L be the set of elements of</u> $PG(m - 1,s)$ <u>having exactly one or two coordinates not equal to zero, then the number</u> $H[m-1,k-1,s,]$ of $(k-1)$ - <u>flats in</u> $PG(m-1,s)$ <u>incident with a point of L is given by:</u>

$$H[m-1,k-1,s] = \phi[m-1,k-1,s] - \theta[m-1,k-1,s] \qquad (2.9)$$

$$\text{where} \quad \theta[m-1,k-1,s] = \frac{\prod_{i=0}^{m-1}[(s^{m-k}-1) - (s-1)i]}{s^{(m-k)(m-k-1)/2}\prod_{i=0}^{m-k-1}(s^{m-k-i}-1)} \qquad (2.10)$$

PROOF: This theorem has been proved by Dowling [1970] in a more general setting and he will publish the results shortly in a separate paper. However, note that the theorem as stated above is just a restatement of a combinatorial confounding problem in terms of the finite projective geometry PG(m - 1,s).

From theorem 2.2 and theorem 2.3 we then have the following corollary:

CORROLARY 2.1. The number N[m-1, k-1, s, 0] of singular saturated main effect plans of type D is given by:

$$N[m-1,\ k-1,\ s,\ 0] = H[m-1,\ k-1,\ s] \cdot s^k \cdot \binom{s^{m-k}}{m(s-1)+1} \qquad (2.11)$$

Next, we claim the following theorem:

THEOREM 2.4. N[m-1, k-1, s, 0] is a lowerbound to T[m,s,0].

PROOF: It is sufficient to exhibit for one particular m and s a set of m(s - 1) + 1 treatment combinations which is not of type of D and which leads to a singular saturated main effect plan. Consider the plan: {(0000), (0011), (0101), (1111), (1010)} for the 2^4 factorial. The matrix X_{11} is then

$$X_{11} = \begin{array}{c} \begin{array}{ccccc} \mu & A/2 & B/2 & C/2 & D/2 \end{array} \\ \begin{bmatrix} 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & -1 \end{bmatrix} \end{array} \qquad (2.12)$$

By inspection of the plan we see that no main effect nor any two factor interaction is completely confounded with the mean. Also it can be verified easily that $|X_{11}| = 0$. Finally, in some cases $N[m-1, k-1, s, 0]$ is equal to $T[m, s, 0]$, e.g. in the 2^3 factorial. Hence we have the conclusion:

$$N[m-1, k-1, s, 0] \leq T[m, s, 0]. \tag{2.13}$$

3. DISCUSSION

The determination of $T[m, s, 0]$ is not so easy and this problem belongs to a more general problem of determining the values which the determinant $|X_{11}|$ can assume when arbitrary selections of $m(s - 1) + 1$ treatment combinations are made from among s^m treatment combinations. The next problem is then to find exactly how many selections belong to a particular determinent. There are of course various ways in which these problems can be solved and methods of attacks are being studied *in extenso*.

4. LITERATURE CITED

[1] Banerjee, K. S. and Federer, W. T. "On estimation and construction in fractional replication". Annals of Mathematical Statistics, Vol. 37 (1966), pp 1033-1039.

[2] Dowling, Th. A. (1970). Private communication.

[3] Kempthorne, O. "The design and analysis of experiments," (1952), Wiley, New York.

[4] Mann, ~~M~~ H. B. "Analysis and design of experiments" (1949), Dover Publications, Inc., New York.